US005091964A

United States Patent [19]

Shimomura

[11] Patent Number: 5,091,964

[45] Date of Patent: Feb. 25, 1992

[54] **APPARATUS FOR EXTRACTING A TEXT REGION IN A DOCUMENT IMAGE**

[75] Inventor: **Shoji Shimomura**, Tokyo, Japan

[73] Assignees: **Fuji Electric Co., Ltd.**, Kawasaki; **Fujifacom Corporation**, Hino, both of Japan

[21] Appl. No.: **678,217**

[22] Filed: **Apr. 1, 1991**

[30] **Foreign Application Priority Data**

Apr. 6, 1990 [JP] Japan 2-91654

[51] **Int. Cl.**[5] **G06K 9/34**

[52] **U.S. Cl.** **382/9**; 382/18; 382/48; 382/51; 358/462; 358/464

[58] **Field of Search** 382/9, 18, 43, 48, 50, 382/51; 358/456, 460, 462, 432

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,403,257 9/1983 Hsieh 358/462
4,447,830 5/1984 Stoffel 358/462
4,741,046 4/1988 Matsunawa et al. 358/464

FOREIGN PATENT DOCUMENTS

344788 2/1991 Japan 606 X/9/20

Primary Examiner—David K. Moore
Assistant Examiner—Jose L. Couso
Attorney, Agent, or Firm—Spencer & Frank

[57] **ABSTRACT**

A peripheral distribution of filled pixels in a document image is calculated by projecting the filled pixels in an X-axis or a Y-axis direction. A bottom part in the peripheral distribution is detected. The document image is divided into a plurality of primary image regions in accordance with a dividing line intersecting the bottom part in the X-axis or Y-axis direction, so that the document image is classified into text regions, drawing regions and picture regions. Thus, the text regions can be extracted automatically from the document image without requiring a specific manual operation for extracting text regions by an operator.

14 Claims, 7 Drawing Sheets